United States
Environmental Protection
Agency

Air And Radiation
(6604J)

EPA 402-R-93-001B
January 1993

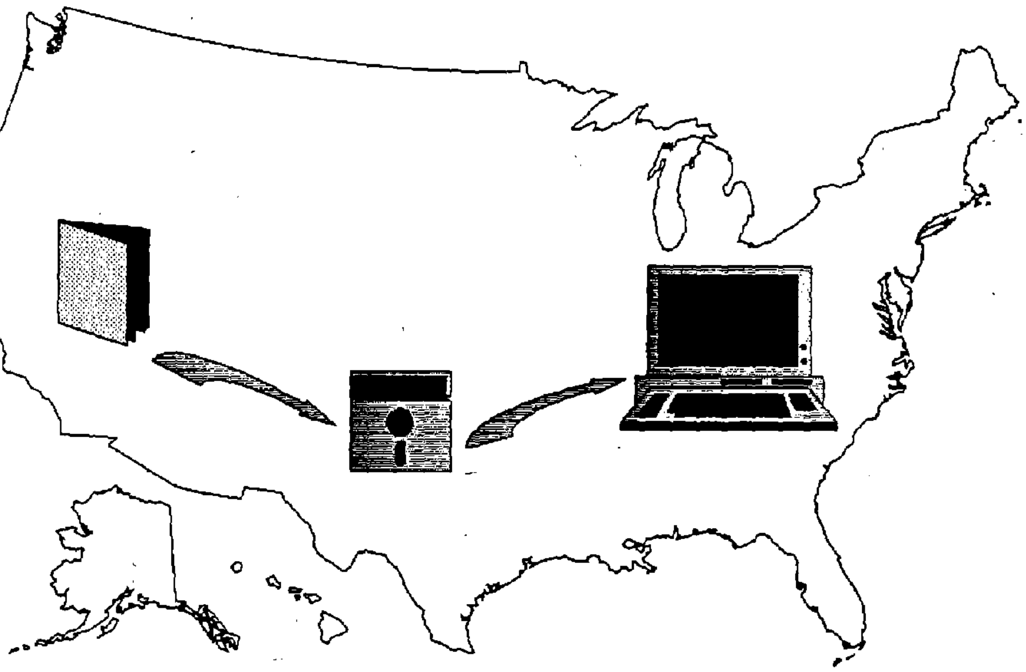
PB93-134799

EPA

National Radon Database

Volume 2: State/EPA Residential Radon Survey

PB93-134799



PB93-134799

NATIONAL RADON DATABASE DOCUMENTATION
Volume 2

The EPA/State Residential Radon Surveys: Year 2

U.S. Environmental Protection Agency
Washington, D.C. 20460

Sharon White
Work Assignment Manager

January 1993

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

1. Introduction

The National Radon Database has been developed by the U.S. Environmental Protection Agency (EPA) to distribute information collected in two recently completed radon surveys:

1. The EPA/State Residential Radon Surveys, Years 1 to 6; and
2. The National Residential Radon Survey.

The State Residential Radon Surveys were conducted in 42 states and 6 Indian lands to characterize the state-wide distribution of radon screening measurements in the lowest livable area of owner-occupied homes. The National Residential Radon Survey was designed to provide an estimate of the national frequency distribution of annual average radon concentrations in occupied residences. Data and documentation for each survey are available through the National Technical Information Service (NTIS).

1.1 GOALS OF THE EPA/STATE RESIDENTIAL RADON SURVEYS

These surveys are statistically valid at the state level and regional levels within each state. The results represent screening measurements and should not be used to estimate annual averages or health risks. Although states and portions of states have been characterized with high or low indoor radon results, the only way to determine the indoor radon level of an individual house is to test. EPA recommends that all homes test for elevated indoor radon levels.

In response to the growing concern about potential health risks associated with indoor radon exposure, the EPA initiated a program in 1986 to assist states in measuring radon concentrations in homes. The importance of this program was confirmed by the Indoor Radon Abatement Act of 1988, Section 305, which directed the EPA to provide technical assistance to the States in assessing radon concentrations in homes. Through this program, the EPA provided assistance to states in the selection and testing of a probability-based sample of houses. Research Triangle Institute (RTI) supported EPA and the states in this effort during the six years of surveys. Assistance was provided in survey design, interviewer training, sample selection, data processing, and data analysis. In addition, the Agency provided the charcoal canisters used in the surveys and also provided all laboratory analysis.

The goals of the state radon surveys were twofold. Some measure of the distribution of radon levels among residences was desired for major geographic areas within each state and for each state as a whole. In addition, it was desired that each state survey would be able to identify areas of potentially high residential radon concentrations ("hot spots") in the state, enabling the state to focus its attention on areas where indoor radon concentrations might pose a greater health threat.

To ensure the discovery of elevated radon concentrations within a home, the charcoal canisters were exposed under closed-house conditions during the winter and were placed on the lowest livable level. Thus, the estimates of indoor radon concentration provided by the surveys reflect a worst-case scenario and maximize the likelihood of identifying residences with high radon concentrations. The screening measurement provides a measurement of the maximum concentration to which occupants may be exposed. A screening measurement also provides a basis for determining whether additional measurements are needed for making a mitigation decision. Data from these state surveys should not, however, be used directly in assessing health risks, because the screening measurements may overstate annual average concentrations in living areas of these homes.

Since the winter of 1986-87, the EPA has assisted 42 states in conducting surveys of indoor ^{222}Rn concentrations. The 42 states and 6 Indian lands radon surveys included in the National Radon Database were carried out during the six years of the program as listed in Table 1-1. Probability-based surveys also were conducted in six selected Indian lands during four of the six years of the program. The use of probabilities in making house selections allows the results to be extrapolated beyond the sample itself to a well-defined population of homes through the use of sampling weights, which are included in the database for all surveys except Colorado and Connecticut.[1] The sampling weights should be used as described in this documentation to replicate the population estimates presented here. In addition, sample data from state surveys conducted by Colorado and Connecticut are included in the Year 1 database. The sampling weights for these states are set to a value of 0 in the database.

A two-day deployment of open-faced charcoal canisters was used by 24 states and 3 Indian lands during the first three years of the state radon survey assistance program. During these years, a diffusion barrier charcoal canister was developed specifically to be less sensitive to the effects of humidity and air flow than the open-faced canister. Two-day deployment of barrier canisters was used by the eight states and two Indian lands in Year 4 of the program. The exposure period for the barrier canisters was increased from two days to seven days for Years 5 and 6. All devices were analyzed promptly at the EPA laboratory in Montgomery, Alabama. Estimates of the relative measurement error as a percentage of the measured concentration were provided by the laboratory and are included in the database. The performance of the charcoal canisters was monitored periodically through the use of unexposed canisters, canisters exposed to known levels of ^{222}Rn, and collocated canisters.

The database now contains data on short-term screening measurements made on the lowest livable level of over 63,000 randomly selected houses during the winter heating season. Survey results for the 42 states and 6 Indian lands are listed in Table 1-2, which

[1] Colorado and Connecticut conducted state surveys and these data are included in the database for Year 1. Because sampling weights could not be determined for these samples, the survey results for these two states should not be extrapolated beyond the sample. The States of Delaware, Florida, New Hampshire, New Jersey, New York and Utah also have conducted their own surveys. Information concerning these state surveys is included in Appendix D.

shows for each state and Indian land the number of homes tested, the estimated number of residences in the target population, population estimates of the arithmetic mean (average) screening measurement radon concentration, and the estimated population percentage of homes with screening measurements over 4 pCi/L and over 20 pCi/L. Due to the lack of sampling weights for Colorado and Connecticut, reported results are applicable only to the sample households. Results are reported separately for the six Indian lands included in the database.

The geographical distribution of estimated mean screening-level radon concentrations is depicted in Figures 1-1 and 1-2 for the 38 states in the contiguous U.S. with probability-based survey results. These states contain 225 sub-state regions. In Figure 1-1 the regions are grouped into three categories using the estimated regional mean screening measurement: 0 to 2 pCi/L; 2 to 4 pCi/L; and greater than 4 pCi/L. In Figure 1-2, the top 60 regions with an estimated mean screening level over 4 pCi/L are displayed in three more-detailed categories: 4 to 6 pCi/L; 6 to 8 pCi/L; and greater than 8 pCi/L.

Figure 1-3 shows a map of the 10 EPA regions used to define the target population for the surveys of Indian lands. The names and addresses of the EPA regional office radon contacts are included in Appendix D.

Table 1-1 Summary of Six Years of the EPA/State Residential Radon Surveys

Year 1, 1986-87 heating season: ten states

Alabama	(AL)	Michigan	(MI)
Colorado	(CO)	Rhode Island	(RI)
Connecticut	(CT)	Tennessee	(TN)
Kansas	(KS)	Wisconsin	(WI)
Kentucky	(KY)	Wyoming	(WY)

Year 2, 1987-88 heating season: seven states and one Indian land

Arizona	(AZ)	Minnesota	(MN)
Indiana	(IN)	Missouri	(MO)
Massachusetts	(MA)	North Dakota	(ND)
Region 5 Indian Land	(R5)	Pennsylvania	(PA)

Year 3, 1988-89 heating season: eight states and two Indian lands

Alaska	(AK)	New Mexico	(NM)
Georgia	(GA)	Ohio	(OH)
Iowa	(IA)	Vermont	(VT)
Maine	(ME)	West Virginia	(WV)
Region 6 Indian Land	(R6)	Region 7 Indian Land	(R7)

Year 4, 1989-90 heating season: nine states and two Indian lands

California	(CA)	Nevada	(NV)
Hawaii	(HI)	North Carolina	(NC)
Idaho	(ID)	Oklahoma	(OK)
Louisiana	(LA)	South Carolina	(SC)
Nebraska	(NE)	Navajo Nation	(RN)
Billings, MT IHS Area	(RB)		

Year 5, 1990-91 heating season: six states and one Indian land

Arkansas	(AR)	Mississippi	(MS)
Illinois	(IL)	Texas	(TX)
Maryland	(MD)	Washington	(WA)
Eastern Cherokee Nation	(RC)		

Year 6, 1991-92 heating season: two states

Montana	(MT)	Virginia	(VA)

Table 1-2 EPA/State Residential Radon Survey Results, Years 1 to 6

State/Indian Land	# Homes Tested	Estimated # Homes in Population	Screening-Level Estimates Arithmetic Mean	Percent > 4 pCi/L	Percent > 20 pCi/L
AK	1,127	38,287	1.7	7.7	0.6
AL	1,180	565,603	1.8	6.4	0.3
AR	1,535	411,395	1.2	5.0	0.3
AZ	1,507	481,861	1.6	6.5	0.1
CA	1,885	2,232,780	1.0	2.4	0.1
CO*	1,443	1,443	5.2	41.5	2.7
CT*	1,451	1,451	2.8	18.5	0.9
GA	1,534	826,452	1.8	7.5	0.0
HI	523	67,044	0.2	0.4	0.0
IA	1,381	593,815	8.9	71.0	7.5
ID	1,266	187,124	3.3	20.3	1.1
IL	1,450	1,537,325	2.9	19.2	0.8
IN	1,914	992,634	3.7	28.5	1.5
KS	2,009	509,496	3.1	22.5	0.7
KY	879	585,655	2.7	17.1	1.5
LA	1,314	432,162	0.5	0.8	0.0
MA	1,659	1,010,301	3.4	22.7	1.3
MD	1,126	761,456	3.1	18.9	1.4
ME	839	236,917	4.1	29.9	1.9
MI	1,989	1,519,962	2.1	11.7	0.4
MN	919	966,496	4.8	45.4	1.4
MO	1,859	998,706	2.6	17.0	0.7
MS	960	352,285	0.9	2.2	0.1
MT	833	151,605	6.0	42.2	4.7
NC	1,290	1,114,747	1.4	6.7	0.3
ND	1,596	194,315	7.0	60.7	4.3
NE	2,027	310,857	5.5	53.5	1.9
NM	1,885	191,090	3.2	21.8	0.8
NV	1,562	93,004	2.0	10.2	0.8
OH	1,734	1,843,743	4.3	29.0	2.8
OK	1,637	538,309	1.1	3.3	0.0
PA	2,389	2,262,234	7.7	40.5	7.9
RI	376	165,646	3.2	20.6	1.9
SC	1,089	505,281	1.1	3.7	0.3
TN	1,773	741,551	2.7	15.8	1.3
TX	2,680	2,216,326	1.0	3.6	0.2
VA	1,156	972,708	2.3	13.9	1.2
VT	710	117,523	2.5	15.9	0.9
WA	1,935	711,965	1.7	8.8	1.3
WI	1,191	933,700	3.4	26.6	0.8
WV	1,006	324,038	2.6	15.7	0.8
WY	777	74,234	3.6	26.2	1.8
SUBTOTAL	59,395	28,773,526			
R5	934	5,328	2.9	19.7	1.3
R6	740	5,443	2.7	16.9	0.8
R7	669	8,478	5.4	34.9	2.7
RB	187	5,834	2.9	22.3	0.0
RC	594	786	0.8	1.7	0.0
RN	772	33,354	1.7	8.3	0.0
SUBTOTAL	3,896	59,223			
TOTAL	63,291				

(*) - Colorado and Connecticut results apply only to those homes tested in the survey.

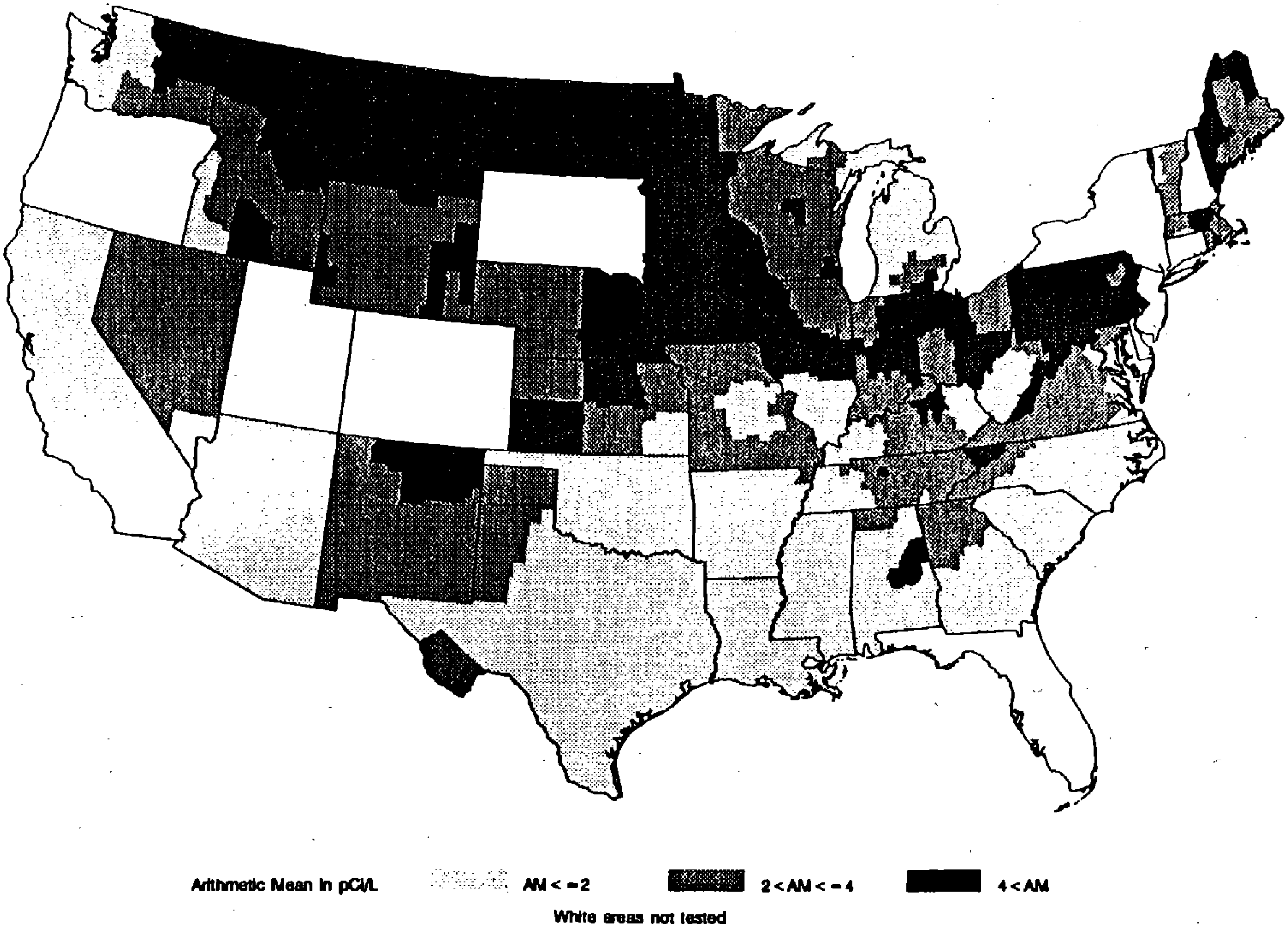


Figure 1. Distribution of Arithmetic Means of Screening Measurements in 225 Regions

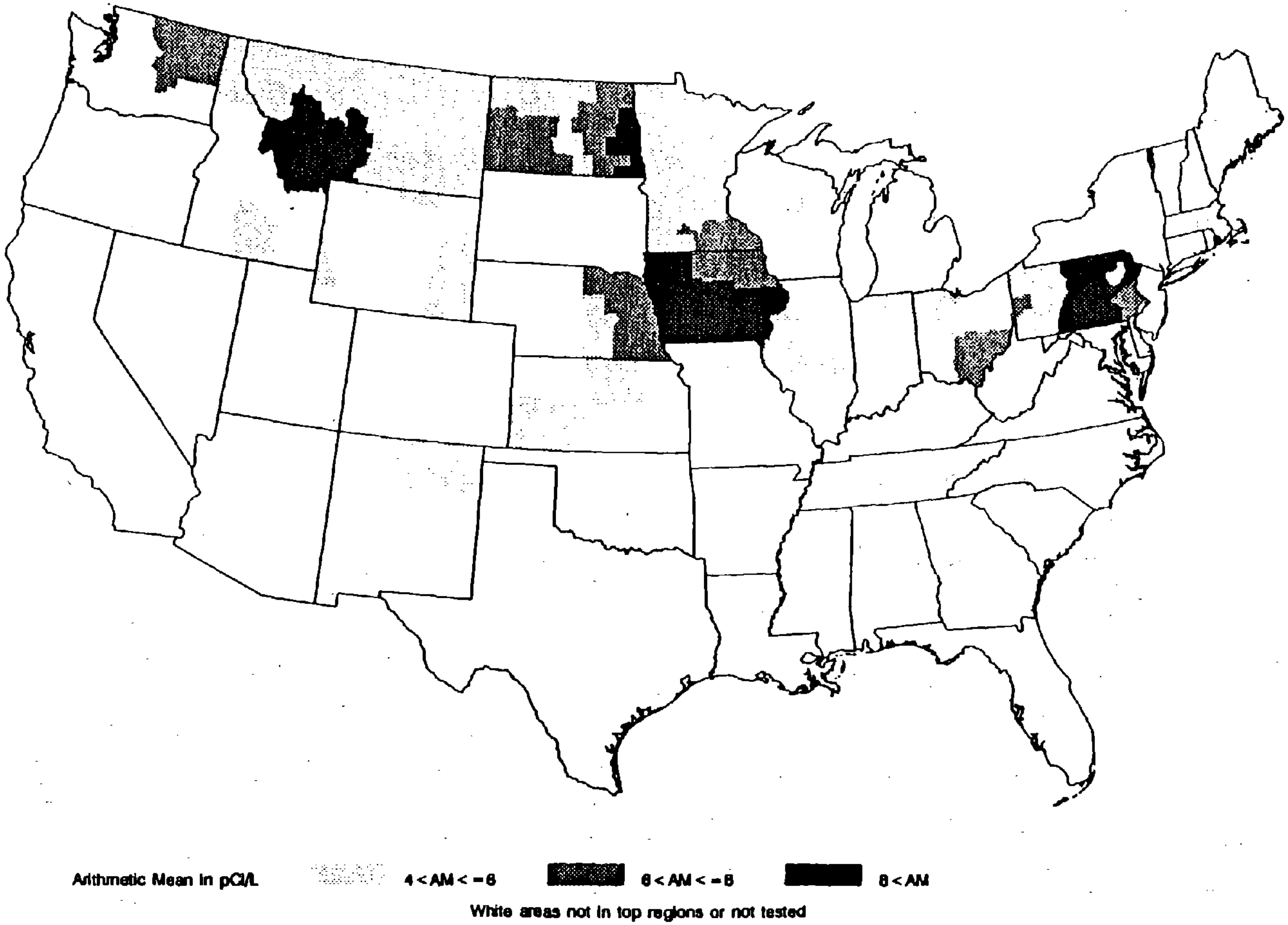


Figure 2. Distribution of Arithmetic Means of Screening Measurements in the Top 60 Regions

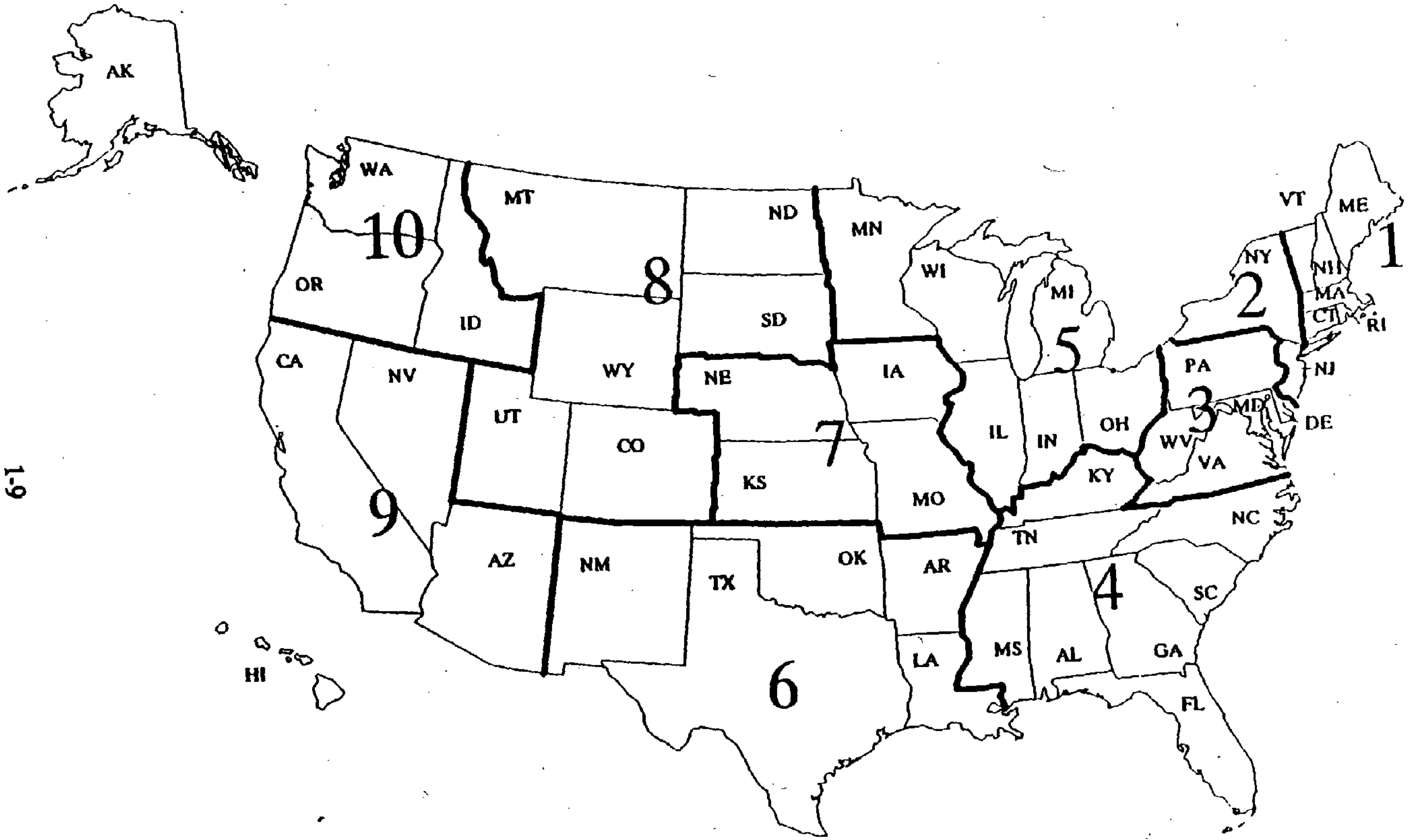


Figure 3. EPA Regions

1.2 SUMMARY OF THE YEAR 2 SURVEYS

During the winter and spring of 1987-88, EPA assisted seven states in conducting state-wide radon surveys and in addition assisted the Indian Health Service (IHS) in conducting a survey of homes on Indian lands located in EPA region 5. The Year 2 states are:

Arizona	(AZ)	Missouri	(MO)
Indiana	(IN)	North Dakota	(ND)
Massachusetts	(MA)	Pennsylvania	(PA)
Minnesota	(MN)	EPA Region 5 Indian Lands	(R5)

For each of the seven states conducting surveys during Year 2, a random sample of residences with listed telephone numbers was selected. For the survey of Indian land, a probability sample of residences was selected for the survey from a listing of all residences located on Indian lands in Michigan, Minnesota, and Wisconsin. Although the sample for Indian lands was selected without regard to the existence of a listed telephone number, information on telephone status was obtained from those selected into the sample.

For each of the Year 2 states, the sample for the state radon survey was a stratified random sample of directory-listed telephone numbers. Geologic groupings were then used as strata for sample selection purposes. A cooperative effort between the state, EPA, and the U.S. Geological Survey (USGS) geologists resulted in the ranking of each of these geologic regions according to the geologists' predictions of the number of homes with high radon concentrations that would likely be found in those areas. This permitted some oversampling of homes in areas where radon levels were expected to be higher. For convenience in selecting the sample of telephone numbers, county boundaries were used to delineate the geologic regions.

The homes to receive measurements were selected as follows. First, a probability sample of residential telephone numbers was selected from a sampling frame constructed from the telephone directories for all communities in the state. Telephone numbers in the designated higher radon strata were sampled at higher rates. After the sample was selected, it was partitioned into sample waves, each consisting of a random subsample of 50 telephone numbers. The sequentially numbered waves were implemented in a specified numerical order, permitting the generation of statistical estimates for the random subpart of the sample represented by the implemented waves.

Proceeding sequentially from wave to wave, telephone calls were made to the sample residential telephone numbers. The interviewer first screened for survey eligibility, which required that the dwelling have a floor on or below grade level and, for reasons of liability, that it be owner-occupied. Once survey eligibility was established, the owner-occupant was requested to participate in the survey. Descriptive material about radon and about the survey was provided either before or after solicitation of cooperation. Those agreeing to participate were provided with a canister and instructions for its use, either by mail person. Participants, after exposing the canister for 48 hours, sent together with a short questionnaire describing where and when the readings had been taken to the EPA Laboratory in Alabama.

The state radon screening survey results are statistically valid at the state and sub-state regional level. The assignment of counties to regions within each state is detailed in Table C-1 of Appendix C. The number of radon detectors (charcoal canisters) also is shown for each county in this table. Table 1-3 contains population estimates for selected parameters of the regional and state-wide radon distribution. These estimates were obtained using the appropriate sampling weights, as described in Section 3.3. The table contains estimates of the mean (average) screening measurement, the median, the geometric mean, the 75th and 90th percentiles, and the percent of houses over 4 pCi/L and over 20 pCi/L.

Table 1-3 Parameter Estimates from the Distribution of Indoor Radon Screening Measurements in Year 2 Surveys, by State and Region (1987-88)

	Number Houses Tested	Est. No. Houses in Population	Arith. Mean pCi/L	Geo. Mean pCi/L	Median pCi/L	75th Percentile pCi/L	90th Percentile pCi/L	% Houses > 4 pCi/L	% Houses > 20 pCi/L
					Arizona				
State	1,507	481,861	1.6	1.0	1.1	2.0	3.4	6.5	0.1
Region 1	161	21,832	1.7	1.0	1.0	2.1	3.8	7.9	0.0
Region 2	200	40,789	1.1	0.7	0.8	1.6	2.4	1.2	0.0
Region 3	1,146	419,240	1.6	1.0	1.1	2.0	3.4	6.9	0.1
					Indiana				
State	1,914	992,634	3.7	2.1	2.2	4.5	8.2	28.5	1.5
Region 1	489	202,242	3.1	1.7	1.7	3.3	6.7	19.4	1.2
Region 2	456	171,420	4.1	2.6	2.8	5.0	9.1	36.2	1.1
Region 3	448	382,242	4.2	2.4	2.6	5.2	8.4	33.7	1.9
Region 4	313	97,958	3.8	2.1	2.1	4.3	7.8	27.7	1.9
Region 5	208	138,773	2.8	1.6	1.7	3.2	5.6	18.5	1.1
					Massachusetts				
State	1,659	1,010,301	3.4	1.9	1.9	3.8	7.0	22.7	1.3
Region 1	103	63,657	2.2	1.4	1.5	2.7	4.4	14.7	0.0
Region 2	81	49,424	2.8	1.7	1.6	3.3	6.3	22.2	0.0
Region 3	47	28,756	3.2	1.7	1.6	3.8	8.9	20.9	0.0
Region 4	125	76,507	1.9	1.3	1.3	2.1	4.0	9.6	0.8
Region 5	219	134,505	4.6	2.8	2.8	5.8	10.4	37.5	2.7
Region 6	391	235,925	4.1	2.3	2.2	4.0	7.8	25.2	2.6
Region 7	201	122,397	4.0	2.4	2.6	4.6	7.8	32.8	0.9
Region 8	162	97,906	3.0	1.8	1.9	3.6	6.4	20.4	0.6
Region 9	114	68,234	2.8	1.7	1.8	3.5	5.9	21.2	0.9
Region 10	141	86,787	3.4	1.3	1.4	2.8	4.1	11.5	0.7
Region 11	75	46,204	1.7	1.3	1.3	2.5	3.0	2.8	0.0
					Minnesota				
State	919	966,496	4.8	3.5	3.7	6.0	9.5	45.4	1.4
Region 1	127	57,880	3.0	2.1	1.8	3.3	5.4	17.5	0.8
Region 2	160	136,070	4.6	3.3	3.4	5.5	9.0	41.7	1.3
Region 3	298	427,818	4.2	3.3	3.3	5.4	7.7	40.1	1.2
Region 4	144	140,431	6.3	4.6	4.9	8.4	10.9	62.3	2.1
Region 5	190	204,296	5.7	4.1	4.4	7.5	11.3	55.4	1.5

Table 1-3 Parameter Estimates from the Distribution of Indoor Radon Screening Measurements in Year 2 Surveys, by State and Region (1987-88) (Continued)

	Number Houses Tested	Est. No. Houses in Population	Arith. Mean pCi/L	Geo. Mean pCi/L	Median pCi/L	75th Percentile pCi/L	90th Percentile pCi/L	% Houses > 4 pCi/L	% Houses > 20 pCi/L
					Missouri				
State	1,859	998,706	2.6	1.6	1.6	3.0	5.4	17.0	0.7
Region 1	624	237,874	3.8	2.5	2.5	4.6	8.5	29.6	1.5
Region 2	90	55,488	2.4	1.5	1.4	2.2	5.0	16.6	0.0
Region 3	218	110,679	1.9	1.1	1.0	2.3	4.1	11.1	0.4
Region 4	437	344,238	2.3	1.7	1.6	2.6	4.5	13.6	0.2
Region 5	262	134,565	2.3	1.4	1.3	2.5	4.3	12.1	0.7
Region 6	228	115,863	2.4	1.3	1.2	2.5	5.2	13.0	0.8
					North Dakota				
State	1,596	194,315	7.0	4.8	4.8	8.0	13.3	60.7	4.3
Region 1	423	23,366	7.8	5.2	5.2	8.8	14.4	62.4	5.7
Region 2	121	12,702	6.6	5.1	5.0	7.8	11.4	65.3	4.6
Region 3	131	14,934	5.1	3.7	3.9	5.8	9.0	46.1	1.4
Region 4	470	64,017	8.9	6.2	6.5	10.9	17.5	72.4	7.0
Region 5	241	47,125	4.8	3.8	3.8	5.5	8.3	47.4	1.5
Region 6	210	32,171	6.9	4.3	4.6	7.3	12.3	60.8	2.9
					Pennsylvania				
State	2,389	2,262,234	7.7	3.3	3.0	7.4	16.6	40.5	7.9
Region 1	268	575,397	7.3	3.4	3.3	7.6	16.6	44.2	8.6
Region 2	258	274,099	17.8	8.5	8.9	17.2	37.1	74.0	21.3
Region 3	270	286,598	10.8	4.5	4.2	9.8	24.1	51.3	10.3
Region 4	273	90,944	7.3	3.8	3.5	7.4	19.5	43.9	10.1
Region 5	249	264,748	4.2	2.5	2.1	4.9	9.9	29.4	2.8
Region 6	279	178,362	5.6	2.3	2.0	4.6	10.4	27.5	4.0
Region 7	207	70,173	8.3	3.1	2.7	7.1	16.6	40.4	8.5
Region 8	199	112,551	3.2	2.2	2.0	3.8	7.1	22.8	0.8
Region 9	261	276,968	4.7	2.4	2.3	4.6	8.8	27.4	3.1
Region 10	125	132,394	2.3	1.4	1.5	2.4	4.3	11.3	0.8
					Region 5 Indian Lands				
ALL	934	5,328	2.9	1.6	1.8	3.3	6.3	19.7	1.3
Region 1	204	1,012	1.1	0.6	0.6	1.2	2.3	2.1	0.2
Region 2	269	2,044	3.0	2.0	2.1	3.2	6.3	20.8	1.4
Region 3	183	630	6.1	4.1	4.2	6.4	9.8	54.1	4.4
Region 4	278	1,642	2.7	1.6	1.7	3.0	5.7	16.1	0.8

In summary, each state radon survey is designed to provide statistical estimates of radon concentration

- In owner-occupied residences,
- With listed telephones numbers, and
- A floor at or below ground level.

2. The Sample Design

2.1 THE OVERALL SAMPLING PLAN

The sampling plan for the state radon surveys called for the selection of probability samples of residences in each state. A probability sample is one in which every element in the population has a positive chance of selection, and, for every element in the sample, the selection probability or relative probability is known. Probability sampling permits the extrapolation of survey results to the entire population and, in addition, can permit the calculation of measures of precision for the estimates. Because one of the goals of each state radon survey was the generation of estimates of distributions of residential radon levels for eligible residences in the state as a whole and for the major geographic areas within the state, use of probability sampling was imperative. Probability-based surveys were also necessary to validly compare results from one state with results from another.

2.2 POPULATION DEFINITION AND SAMPLING FRAMES

The target population for the surveys in all seven of the Year 2 states consisted of owner-occupied residences with a permanent foundation and at least one floor at or below ground level and with a telephone number published in the latest directory. The statistical estimates generated from the survey data apply to this population.

In reality, the totality of occupied residences in the state constituted the population of interest. However, as is often the case in survey research, surveying this population was not deemed feasible, for several reasons. First, it was considered inadvisable from a legal point of view to include rental dwellings without first obtaining the permission of the owner. Although procedures could be devised to obtain such permission, the cost of doing so, both in dollars and in delay in the survey schedule, was deemed impractical. Second, homes that had no floor on or below ground level were excluded from the survey target population. Although these homes are no doubt usually rental apartment units, the category would include some owner-occupied condominiums. These were excluded from the target population because radon levels on upper floors were expected to be low, and it was felt that the focus of the survey should be on residences that were potentially at risk. Third, the survey target population was restricted to homes with listed telephone numbers, basically because of time and cost considerations. Sampling of homes without regard to the existence of a telephone would call for an area probability procedure, which requires onsite staff for both listing and data collection and is both expensive and time consuming. The telephone survey approach was used because it offered a more economically feasible alternative. Telephone surveys can be implemented using a relatively small staff working in a central location, and they can be carried out on short notice and within a restricted time schedule.

Two types of samples are commonly used for telephone surveys: random digit dialing samples, for which every possible telephone number is given a positive chance of being selected into the sample, and telephone directory samples, for which only listed telephone numbers are given a chance of selection. In Year 2, each state was given the choice of these two telephone survey methods, and each chose the procedure calling for the selection of listed telephone numbers. There were two major incentives for making this choice. First, the labor involved in telephoning is much less using listed telephone numbers than it is using random digit dialing because the vast majority of listed numbers will be working residential numbers, as compared to only about 20 percent for the random digit dialing technique. Second, names and addresses are available for directory-selected telephone numbers, making possible a mailing of material describing the health risks associated with radon exposure and describing the survey. This second reason was an important consideration for those states wishing to do a mailing prior to the telephone contact.

Two organizations constructed files of listed telephone numbers: Survey Sampling and Donnelley Marketing. While both organizations had comparable sampling frames, Survey Sampling was more restrictive in the selection procedures that they were willing to implement. Only Donnelley Marketing was willing to follow precisely the sample selection procedures developed for the state radon surveys; therefore, samples for six of the seven Year 2 states were purchased from that company.

The North Dakota sample was selected directly from telephone books, rather than from the Donnelley file, but it employed selection procedures very similar to those used for the other six states. The sample for the Indian lands survey was selected from a frame provided by the IHS using procedures described in Section 2.7.

2.3 STRATIFICATION AND SAMPLE ALLOCATION

To improve the precision of the survey estimates, the sampling frame for each of the six states using the Donnelley frame was stratified prior to sample selection. To the extent that the variable(s) used for stratification are correlated with the variable being estimated, the sampling error of the survey estimates can be reduced. The major stratification variable was, therefore, the classification of counties according to the likelihood of finding high residential radon readings in them. The counties within a state were typically classified into three to five groups by the state geologist with assistance from geologists at EPA and USGS. Using the groups provided, the total number of canisters that were expected from each county was estimated, given the total sample size that was agreed upon by EPA and the state. The estimation procedure involved simply applying a sampling rate to the Market Statistics' estimate of housing units for the county for 1988, assuming uniform eligibility and response rates across strata. (Market Statistics, Inc., produces county-level housing unit projections for the nation.)

Some investigation was done of the effects of sampling residences in higher radon strata at higher rates than those in lower radon strata. Use of differential sampling rates could increase the precision of other estimates because of the associated unequal weighting. As a result of this investigation, the sampling rates were typically set with an approximate four to one ratio of the highest to lowest sampling rate. This provided the desired increase in precision of radon level estimates, if the classification carried out by the geologists held. If such classification failed to partition the state into groups that were different on radon level, but instead partitioned it into, for example, groups that were identical with respect to radon levels, the precision of the estimates would not be greatly decreased. The unequal weighting effect (DEFF) associated with unequal sample selection rates was computed for each alternative sample allocation, with the aim of keeping it under about 1.4 or 1.5. This meant that, for the design used, the error variance of characteristics that were uniformly distributed across strata would be no more than 1.4 or 1.5 times that which would have resulted from an equal probability design. (Note that an error variance 1.4 to 1.5 times as large means a sampling error only about 1.2 times as large.)

Once the basic allocation was set, some adjustments of the strata were considered. If, for example, there were some counties for which the expected number of canisters per square mile was extremely small, an alternative allocation was prepared, i.e., moving the county to a stratum with a higher sampling rate. This was done to keep large areas of the state from being covered too sparsely, and thus compromising one of the survey's goals, which was to identify "hot spots" in the state. Additional adjustments of different sorts were carried out. For example, when a large metropolitan area was found to have a very large expected number of canisters, an alternative allocation was done, i.e., assigning it a stratum with a lower sampling rate. This permitted a somewhat higher sampling rate to be used in other portions of the state and improved the likelihood of discovering any "hot spots" that might exist. The alternative sample allocations, together with a description of the advantages and disadvantages of each, were sent to the states. The state then selected the allocation they wished to have implemented, subject to EPA's approval.

The target number of canisters to be placed, a description of the allocation that was chosen by the state, the sampling rates used in the strata, and the expected DEFF for variables that are uniformly distributed across strata are presented for each state it Appendix C.

Following guidelines determined by the agreed upon allocation, the samples for the six states were selected from the Donnelley Marketing files. In all cases, detailed instructions for ordering the file and selecting the sample for each state were prepared. The instructions called for ordering the residential telephone listings in each stratum by the size rank of the county in which the residence was located, then by the census block group or enumeration district. The listings were finally ordered by telephone number. This ensured maximum geographic spread when systematic random sample selection procedures were used.

2.4 SAMPLE SELECTION PROCEDURES FOR THE DONNELLEY FILE SELECTIONS

To permit the unbiased estimation of the sampling errors of the survey estimates of radon characteristics for the state and for major geographic subparts of the state, five independent, systematic random samples were selected from each stratum. To do this, RTI provided the sample size to be selected from each stratum for each of the five samples, a list of the counties that made up each stratum, and the specifications for ordering the file within each stratum. The sample selection instructions that were provided are presented in Table 2-1.

The following variables were requested for each sample selection:

1. State code from the Federal Information Processing Standards (FIPs),
2. County FIPS code,
3. Stratum,
4. Area code,
5. Telephone number,
6. Name,
7. Mailing address,
8. ZIP code, and
9. Sample (or replicate) number (1-5).

Table 2-1 Procedures for Selecting the Sample of Telephone Numbers

1. Sort all residential telephone numbers in the state as specified.

2. Determine the number of listings of residential telephone numbers on the file for the stratum. Call this number L.

3. Identify the sample size specified for the stratum and call this number S.

4. Divide L by S to obtain the Selection Interval I.

5. Select 5 different random numbers between (and including) 0.00000001 and I.

6. Successively add I to the first random number to generate S selection numbers. Round up the S selection numbers for the stratum to identify the sample telephone numbers on the ordered list.

7. Repeat step 6 for each of the other 4 random numbers until all 5 random samples of size S have been selected.

8. When this procedure has been implemented for all strata defined for a state, the state's sample selection is complete.

2.5 SAMPLE SELECTION PROCEDURES FOR NORTH DAKOTA

RTI did not select the sample for North Dakota, but did prepare the sample selection procedures that were used. The state had contracted with a survey research group associated with the University of North Dakota for all data collection activities. The university group maintained an up-to-date collection of all of the state's telephone directories and routinely selected samples for the telephone surveys it conducted.

Sampling rates were assigned to counties, which had been classified into strata as described in Section 2.3. The sampling rates were then ascribed to the separate sections of each directory, according to the county associated with the majority of listings in the section. Five systematic samples were selected from each of the newly defined strata comprised of a grand list of telephone numbers formed by concatenation of the alphabetically ordered lists from the residential sections of each directory belonging to that stratum.

2.6 PARTITIONING THE SAMPLES INTO WAVES

Estimating the exact number of sample selections that would be needed in a state survey to be able to place the desired number of canisters was very difficult. Unknown were the exact proportion of selected numbers that would be working residential numbers, the exact proportion of residential numbers that would be associated with survey-eligible residences, or the proportion of eligible residences that would participate in the study. Another very important unknown was when the weather in the state would become so warm that the closed house requirement for canister deployment could not be met, and the survey would have to be discontinued.

There is a commonly used technique for controlling the number of survey participants in situations where many unknowns are involved in estimating the number of sample selections needed. The procedure involves partitioning the sample into a number of random subsamples and implementing only as many of the subsamples as are needed to achieve the desired number of participants. This technique was used in all seven Year 2 states.

A sample sufficiently large for any reasonable set of assumptions was selected as described above. It was then partitioned into random subsamples, or waves, of 50 telephone listings each. The waves were randomly ordered and numbered sequentially, and they were activated in a specified numerical order by the states. Implementation of the sample in random subparts meant that a state did not need to complete all sample waves.

The procedures used in processing the file and partitioning the sample into waves are described below.

1. The sample of 10-digit telephone numbers was checked for duplicates, which were eliminated, and was checked to verify that the proper number of records had been provided for each replicate in each stratum.

2. Five percent of each replicate was randomly designated to receive duplicate canisters.

3. The total number of waves, W, into which the sample was to be partitioned was determined by dividing the number of records on the file by 50.

4. The waves number 1 through W were put in random order and assigned to the first W records of the file. The wave numbers 1 through W were again placed in a random order and assigned to the second W records on the file, etc., until each record had been assigned a Wave number.

5. The records were ordered by wave number and a Case ID number was assigned sequentially.

2.7 SAMPLE SELECTION PROCEDURES FOR THE INDIAN LANDS SURVEY

The IHS carried out a residential radon survey on Indian lands in Minnesota, Michigan, and Wisconsin.

For this survey, a personal interview procedure was used rather than a telephone interview procedure as had been used for the other surveys. The canisters were also placed and retrieved by the field interviewer.

The population for the survey was also somewhat different for the Indian lands survey. All owner-occupied homes, with at least one floor on or below ground level and located on one of the Indian lands survey locations, were eligible for the survey, whether or not the land on which the house was located was owned by the occupant and whether or not there was a listed telephone number linked to the home. However, each respondent was asked if there was a telephone at the home and if the telephone number was listed.

The IHS constructed a sampling frame for each reservation, noting the name and address of each family living on the reservation. For ease in distributing canisters to each of the reservations and for controlling the overall sample size, each reservation, denoted by the subscript h, was assigned a specific sample size, n'_h. The n'_h constituted the expected sample sizes that would provide the desired distribution and total number of sample cases across reservations. The addresses within each reservation were put in a random order, and the first n'_h addresses on the list were assigned to the primary sample. The following $1/2\ n'_h$ addresses on the list were assigned to the secondary sample, which was to be used, in the order specified, as needed.

The interviewer visited all of the n'_h cases in the primary sample, determined survey eligibility, and attempted to place a canister in each eligible home. Some primary sample cases were found to be ineligible for the survey. If, for example, the family had moved from the reservation and left their reservation home vacant, the sample case would be classified as "not survey eligible." In rare occasions, a refusal was obtained for survey-eligible home in the primary sample. Whenever participation was obtained from fewer than n'_h eligible homes in the primary sample, the secondary sample cases were worked in the order assigned until detectors were placed in exactly n'_h eligible homes on the reservation.

3. Estimation Using Survey Results

3.1 CALCULATION OF SAMPLING WEIGHTS

Because most of the states used unequal probability sample designs for their state radon surveys, sampling weights that account for the unequal probabilities of selection must be used to generate unbiased population estimates from the survey data. Sampling weights that reflect only the differential selection probabilities would be adequate if 100 percent response rates and participation rates were achieved. However, this level of response was not obtained. For the state radon surveys, some of the sample cases failed to complete a screening interview, either because they were never successfully contacted or because they refused to provide the screening information. Whether or not they were in fact eligible was, therefore, never determined. For other cases, the screening information was provided, and the housing unit was determined to be eligible for the survey, but a canister reading was not successfully linked to the case. There are numerous reasons why this might have occurred. The canister may not have been read because it was never deployed; it may have been deployed but never returned; or it may have been returned but not received in time to be included in the analysis. In addition, clerical or keying errors associated with matching criteria could have prevented matching canister readings with the proper cases. To compensate for the missing information, a weighting class adjustment was used. This procedure increased the sampling weights of participants to compensate for the missing information from nonparticipants. The steps used in calculating sampling weights and adjustments for the seven Year 2 states are described below.

The first step in calculating the sampling weight was determined from the information provided by Donnelley Market Services (or by the state, for North Dakota). For each stratum in the sample, RTI was provided with the number of listings from which the sample was selected. The number of selections that should be made was specified. Using this information, the first component of the sampling weight was computed for each stratum and used for all selections from that stratum. For any stratum h, the first sampling weight component was calculated as

$$W'_h = N_h / [(5)(n_h)], \qquad (1)$$

because five samples of size n_h were selected from the N_h listings in stratum h.

As was described in Chapter 2, each state's sample was randomly partitioned into waves of 50 listings each, each wave being in effect a probability sample of the entire state. Although all waves were available for use in the state radon survey, not all were used. The second component of the sampling weight represented the portion of the sample waves that were included in the analysis. Any wave for which at least 45 of the 50 cases were completed was considered to have been implemented, and it was referred to as an "active" wave. Computer runs were made on the Control/Screening Form file to determine which waves would be classified as "active" and included in the analysis and which would not. For each state, the sampling weight component was computed reflecting the proportion of waves classified as active. This was merely the total number of waves of 50 listings divided by the number of waves classified as active waves, or V/v. Only cases in the v active waves were used in the remaining calculations and in the analysis.

Next an unadjusted sampling weight was calculated for every selected case in every active wave, regardless of the response or participation status of the case. This weight was merely the product of the two weight components.

$$W''_h = (W'_h)(V/v) \qquad (2)$$

Next, every record in every active wave was compared to the file of canister readings and, by matching on House ID number, was classified as a participant or a nonparticipant. All active wave cases classed as participant would be used in the analysis, because they were in an active wave and had a canister reading. To adjust for missing canister readings for the survey eligible, all active wave nonparticipant cases were further classified according to eligibility status. The following groups were formed for the active wave cases:

Group A:	Participants (all eligible cases for which a canister reading was available).
Group B:	Survey eligible nonparticipants.
Group C:	Nonparticipants, survey eligibility unknown. (All cases for which eligibility information was never obtained.)
Group D:	Nonparticipants known to be ineligible for the survey.

These four groupings were used in calculating the adjustments for nonresponse.

Five weighting classes were formed within each stratum, corresponding to the five replicates used in the sample selection. Within each weighting class, an adjustment-for-nonresponse factor was computed, as follows.

First, an estimate of the proportion of cases that were survey-eligible was computed:

$$A'_{sh} = \frac{|\Sigma W''_{shi}|_A + |\Sigma W''_{shi}|_B}{|\Sigma W''_{shi}|_A + |\Sigma W''_{shi}|_B + |\Sigma W''_{shi}|_D} \quad (3)$$

where

$|\Sigma W''_{shi}|_A$ =	sum of the unadjusted sampling weights over all nonparticipants in the s replicate in stratum h,
B =	survey-eligible nonparticipants, and
D =	nonparticipants known to be ineligible.

The proportion A'_{sh} was used to estimate the proportion eligible among those for whom eligibility has not been determined. This figure was needed to determine the nonresponse adjustment factor for each replicate s within each stratum h:

$$A_{shi} = \frac{\left|\Sigma W''_{shi}\right|_A + \left|\Sigma W''_{shi}\right|_B + A'_{sh}\left|\Sigma W''_{shi}\right|_C}{\left|\Sigma W''_{shi}\right|_A} \qquad (4)$$

where

$\left|\Sigma W''_{shi}\right|_C$ = sum of the unadjusted weights over all nonparticipants with unknown eligibility and where all other terms are as defined above.

The final sampling weight was then calculated for each sample case in every active wave as:

$$W_{shi} = (W''_{shi})(A_{shi}) , \qquad (5)$$

and the sampling weight W_{shi} was used as the sampling weight in all analysis.

3.2 CALCULATING SAMPLING WEIGHTS FOR THE INDIAN LANDS SURVEY

A modification of the above procedures was used for the Indian lands survey. A negative binomial distribution was assumed in which n''_h sample homes were contacted on a reservation to obtain n'_h survey-eligible homes. Some of the n''_h selections came from the primary sample, but some could have come from the secondary sample. The proportion of survey-eligible homes for the reservation was estimated to be:

$$(n'_h-1)/(n''_h-1) \qquad (6)$$

The number of survey-eligible homes on the reservation was estimated to be:

$$N_h = [(n'_h-1)/(n''_h-1)]N'_h \tag{7}$$

where

n'_h = desired sample size from reservation h,

n''_h = number of case that needed to be contacted to discover n'_h survey-eligible residences in reservation h, and

N'_h = number of listings on the reservation h sampling frame.

The final sampling weight was calculated for each of the participants providing a usable detector reading as

$$W_{hi} = N_h/(\#\text{usable participants})_h.$$

3.3 ESTIMATING MEANS AND PROPORTIONS

The analytical results were obtained using SESUDAAN, a computer software program developed by RTI for analyzing survey data with complex error structures. Formulas for estimating means and proportions from the state surveys using this program are shown below. Appendix E contains the formulas for estimating means and proportions for the Indian lands.

Define Y^*_r as the true mean radon level for the r^{th} region or reporting group (r=1,...,R). Y^*_r can be estimated as

$$Y^*_r = \frac{\sum_{h=1}^{H} \sum_{i=1}^{n_h} J_{rhi} W_{hi} Y_{hi}}{\sum_{h=1}^{H} \sum_{i=1}^{n_h} J_{rhi} W_{hi}} \qquad (8)$$

where

Y_{hi} = observed radon measurement for the i^{th} eligible household in stratum h ($i = 1, ..., n_h, h = 1, ..., H$);

W_{hi} = sampling weight associated with Y_{hi}; and

$$J_{rhi} = \begin{cases} 1 & \text{if } i^{th} \text{ eligible household in stratum h is in the } r^{th} \text{ region} \\ 0 & \text{otherwise.} \end{cases}$$

The estimated mean for all regions combined (i.e., the statewide estimate) is given by

$$Y^*_o = \frac{\sum_{h=1}^{H} \sum_{i=1}^{n_h} W_{hi} Y_{hi}}{\sum_{h=1}^{H} \sum_{i=1}^{n_h} W_{hi}} \qquad (9)$$

Similarly, define P^*_r as the true proportion of eligible households in the r^{th} region with radon levels exceeding X pCi/l. P^*_r can be estimated as

$$P^*_r = \frac{\sum_{h=1}^{H} \sum_{i=1}^{n_h} J_{rhi} W_{hi} I_{xhi}}{\sum_{h=1}^{H} \sum_{i=1}^{n_h} J_{rhi} W_{hi}} \qquad (10)$$

where

W_{hi} and J_{rhi} are as previously defined and

$$L_{xhi} = \begin{cases} 1 & \text{if measurement on } i^{th} \text{ eligible household in stratum } h \text{ is greater than X pCi/l} \\ 0 & \text{otherwise.} \end{cases}$$

The estimated proportion for all regions combined (i .e., the statewide estimate) is given by

$$P^{*}_{o} = \frac{\sum_{h=1}^{H} \sum_{i=1}^{n_h} W_{hi} I_{xhi}}{\sum_{h=1}^{H} \sum_{i=1}^{n_h} W_{hi}} . \tag{11}$$

4. Methodological Results

The survey methodology used for the second year if the state/EPA radon survey program was reviewed at five different levels:

- First, the coverage of each state survey was assessed. To do this, four different estimates were compared of the number of owner-occupied single family housing units having a telephone, which was the approximate definition of the survey-eligible population. For each state, the survey estimate of this population size was compared to an estimate based on the 1980 Census counts for the state, to an estimate made using current counts from the Donnelley Marketing Service files from which most of the state samples were selected, and to an estimate based on the Market Statistics' projections.
- Second, the response rate and the participation rate obtained in each of the states were computed. These were simply the ratio of the estimated number of respondents to the estimated number of eligible and the ratio of the estimated number of usable canister readings to the estimated number of eligible.
- Third, the number of cases for which eligibility status was never determined was reviewed.
- Fourth, the Control/Screening Forms that were returned by the states to identify the types of errors that the states made in carrying out the survey were reviewed.
- Fifth, all of the problems that occurred throughout the course of all of the Year 2 state radon surveys were assessed to determine whether modifications were needed in survey procedures.

In the sections that follow, each of these assessments of the state radon survey methodology is discussed.

4.1 COVERAGE

The results of the coverage investigation are presented in Table 4-1. For each of the seven Year 2 states, the number of owner-occupied single family housing units with a telephone was estimated using 1980 decennial census information, using Donnelley file counts, using the Market Statistics' estimates, and using state radon survey results. In constructing these estimates, the percentage of housing units that were owner occupied was available by state, but the percentage of owner-occupied housing units that were single unit structures was available only for the nation as a whole. The national average, showing 94 percent of all owner occupied housing as being single unit structures, was therefore used in the calculations for each of the states. In addition, the nationwide estimate of 97 percent was used for the percentage of owner-occupied single structure housing units having a telephone.

Column 3 of Table 4-1 shows an estimate of the approximate number of survey-eligible housing units using 1980 Census counts, and columns 5 and 9 show comparable estimates made from the Donnelley file counts and Market Statistics' estimates, respectively. The ration of the Donnelley estimates to the Census estimates, shown in column 6, vary from a low of 0.81 for Pennsylvania to a high of 1.03 for Minnesota. Column 7 shows comparable ratios for estimates of survey eligible based on Donnelley file counts to those using Market Statistics' data. These ratios vary from a low of 0.66 for Arizona to a high of 0.95 for Minnesota. The two sets of ratios were calculated to get a very rough indicator of what might be missing by using the Donnelley files as sampling frames, without using a supplementary procedure for picking up housing units not linked to a Donnelley listing, but otherwise survey-eligible. The relatively low ratios for Arizona, Indiana, and Pennsylvania indicate a potential for a sizable noncoverage. The extremely high ratios for North Dakota, which used telephone directories, indicate possible multiple listings of individual telephone numbers.

Column 15 shows the ratio of the number of survey eligible in each state, as estimated from the survey itself, to the estimate made directly from the Donnelley frame counts. This ratio was calculated as a measure of the loss suffered because of movers and possibly because of households being difficult to reach by telephone. Recall that the procedures selected a sample of telephone numbers and the housing units linked to those numbers, regardless of whether the address was the same as was given in the sampling frame. Therefore, housing units of movers were picked up, but not to the degree in which they were lost. When someone moves, their telephone number is typically retired for a period of 6 months to a year, unless it is carried to the new home. Therefore, a good many movers were reached at their new home. Intrastate movers who change telephone numbers and those who move in from another state were lost if the move occurs after the cutoff date for directories on which the Donnelley listing are based. The ratio of survey-estimated survey eligible to Donnelley-estimated survey eligible ranged from a low of 0.92 for Arizona to a high of 1.04 for Pennsylvania, indicating very little loss because of movers.

4.2 RESPONSE RATES

Approximate observed response and participation rates are presented in the bottom two rows of Table 4-2. The percentage of known survey-eligible housing units for which the respondent agreed to place a charcoal canister in the home ranges from a low of 82 percent for Pennsylvania to a high of 99 percent for the Indian lands survey and 95 percent for Minnesota. Participation rates show the percentage of known survey-eligible homes for which a usable canister reading was obtained. These percentages vary from a low of 68 percent for Pennsylvania to a high of 95 percent for the Indian lands survey. The high figure for the latter group represents the success of the personal placement and retrieval procedures used in this survey. The highest participation rate for a state survey was 86 percent for Minnesota.

Although the average response rate for known eligible for the seven state surveys was about 89 percent, the average participation rate was only about 75 percent, a drop of about 14 percentage points. Getting people to return their canisters immediately after exposing them for the designated period might be an aspect of data collection that should be given more emphasis. Minnesota had the highest response and participation rates and the smallest difference between the two rates. Their survey staff routinely recontacted people to whom a canister had been sent, but no reading received, to remind them to deploy their canister and to return it immediately after exposure. Minnesota's high participation rate indicates that such a practice might also help other states increase their participation rates.

4.3 UNKNOWN ELIGIBILITY STATUS

The Year 2 states did an excellent job in returning all Control/ Screening Form for all of their activated waves. This aspect of the data collection process received more emphasis in the Year 2 training because it had been found to be a major problem in Year 1. There does, however, seem to be a large number of "eligibility unknown" cases for several states, particularly Arizona and Indiana. This classification was assigned not only to cases in activated waves for which no screening form was received, but also to cases with repeated ring-no-answer calls and to cases for which a contact was made but the screening interview was not completed to the point where eligibility for the survey could be established. It is extremely important to call on different days of the week and different times of the day in order to maximize the chances of contacting a sample case. This type of calling schedule helps to keep the number of ring-no-answer cases to a minimum, which is important because a large number of "eligibility unknown" cases is a source of potential bias in the survey results.

In generating statistical estimates from the survey data, every sample case in every implemented sample wave must be accounted for, including every case for which a screening form was not returned and every case for which eligibility was not determined.

Although these cases were classified as "eligibility status unknown," they cannot be ignored in the estimation process. Sampling weight calculations included adjustments for:

- That portion of the unknown-eligibility category of nonresponse estimated to be survey eligible, and
- The category of nonresponse due to failure of sample eligible to participate in the survey.

These sampling weight adjustments were made in an attempt to reduce the possible bias caused by missing information for sample cases. However, no adjustment can eliminate the potential for such bias. This can only occur when there are no cases for which eligibility status is unknown and no nonresponse.

4.4 ERRORS MADE IN IMPLEMENTING SURVEYS

The principal difficulties encountered in the Year 2 surveys were:

- Incorrect and inconsistent use of result codes, and
- Delayed shipments of completed Control/Screening Forms.

The most common problem associated with all the states was the misuse of final telephone result codes. Delayed shipments of screening forms compounded the result code problem, delaying effective feedback on interviewer performance and causing delays in data processing activities. Completed Control/Screening Forms should be returned on a timely basis, such as once a week, rather than relying on quantity to determine shipment dates.

Table 4-1 Comparison of Estimates of Survey Eligibles

	1980 Census			1987-88 Donnelley			
State	Number of Occupied Housing Units (1)	Percent Owner-Occupied (2)	Estimated Number of Owner-Occupied Single Family Housing Units with Telephone (3)*	Number of Housing Units with Telephone (4)	Estimated Number of Owner-Occupied Single Family Housing Units with Telephone (5)**	Ratio of Donnelley Estimated Eligibles to Census Est. Eligibles (5) ÷ (3) = (6)	Ratio of Donnelley Est. Eligibles to Market Stat. Est. Eligibles (5) ÷ (9) = (7)
AZ	957,032	68.3	596,001	818,279	525,351	.88	.66
IN	1,927,058	71.7	1,259,829	1,528,632	1,030,267	.82	.78
MA	2,032,717	57.5	1,063,723	1,835,742	992,219	.93	.87
MN	1,445,222	71.7	944,829	1,446,682	975,035	1.03	.95
MO	1,793,399	69.6	1,138,114	1,500,745	981,847	.86	.83
ND+	227,664	68.7	142,610	335,574	216,707	1.52	1.38
PA	4,219,606	69.9	2,689,358	3,316,654	2,179,241	.81	.76

	December 1986 Market Statistics			State Radon Survey Estimates				
State	Number of Occupied Housing Units (8)	Estimated Number of Owner-Occupied Single Family Housing Units with Telephone (9)*	Ratio of Market Statistics Estimated Eligibles to Census Estimate Eligibles (9) ÷ (3) = (10)	Sample Sizes (11)	Estimated Number of Survey-Eligible Housing Units (12)	Ratio of Survey-Estimated Eligibles to Census Estimate Eligibles (12) ÷ (3) = (13)	Ratio of Survey-Estimated Eligibles to Market Statistics (12) ÷ (9) = (14)	Ratio of Survey-Estimated Eligibles to Donnelley Estimate Eligibles (12) ÷ (5) = (15)
AZ	1,283,700	799,436	1.34	1,507	481,861	.81	.60	.92
IN	2,016,900	1,318,570	1.05	1,914	992,634	.79	.75	.96
MA	2,185,200	1,145,668	1.08	1,659	1,010,301	.95	.88	1.02
MN	1,564,300	1,022,678	1.08	919	966,496	1.02	.95	.99
MO	1,870,100	1,186,789	1.04	1,859	998,706	.88	.84	1.02
ND+	250,800+	157,103	1.10	1,596	194,315	1.36	1.24	.90
PA	4,530,800	2,881,324	1.07	2,389	2,262,234	.84	.79	1.04

+ North Dakota used their own sample instead of the Donnelley list.

* Assuming 94 percent of owner-occupied units are one unit structures (1983). Also assuming 97 percent of housing units have a telephone (1981).

** Assuming column (2) percent owner-occupied and that 94 percent of these are one unit structures.

Table 4-2 Disposition of Sample Cases

	AZ	IN	MA	MN	MO	ND	PA	RS
Sample Waves Activated	1-19 21-126	1-12 21-125	1-11 21-110	1-4 21-50	1-11 21-99	1-10 21-90	21-140	--
Sample Waves Used in Analysis	1-14 21-126	1-12 21-22 24-125	1-11 21-106 109-110	1-4 21-50	1-10 21-99	1-10 21-90	21-140	--
C/S Forms Received	6,001	5,836	4,950	1,700	4,428	3,976	6,000	1,444
Case Used in Analysis	6,000	5,800	4,950	1,700	4,450	4,000	6,000	
Status Eligibility Status, Code Canister Acceptance								
A1 Eligible, Accepted	1,954	2,237	2,047	1,018	2,250	1,845	2,905	976
A2 Eligible, Refused	142	310	332	52	341	186	621	9
C Eligibility unknown	2,103	1,729	584	157	513	501	730	33
O Not Eligible	944	863	1,584	318	870	977	1,115	390
D Not a Residence	858	697	403	155	454	467	627	36
Total	6,001	5,836	4,950	1,700	4,428	3,976	6,000	1,444
U Usable Readings	1,509	1,914	1,659	919	1,816	1,596	2,389	934
Response Rate ($A_1/(A_1 + A_2)$)	93.2%	87.8%	86.0%	95.1%	86.8%	90.8%	82.4%	99.0%
Participation Rate ($U/A_1 + A_2$)	72.0%	75.7%	69.7%	85.9%	71.8%	78.6%	67.8%	94.8%

APPENDIX A

Installation Procedures

INSTALLATION PROCEDURES

I. EXTRACTING DATA FROM THE DISKETTE

The diskette you have received contains three files:

- **DATA.FIL** - a compressed version of the screening measurement data collected in one year of the EPA/State Residential Radon surveys.
- **EXTRACT.EXE** - an executable program to extract and store the expanded version of the survey data file on your hard disk. The extract program will run on any IBM-compatible personal computer using the MS-DOS operating system, Version 2.0 or higher.
- **READ_ME.1ST** - a copy of these instructions.

To expand the compressed file onto your hard disk, place the diskette in the appropriate drive and change to this drive. (For example, type **A:** then press the Enter key.) Run the program by typing the command **EXTRACT**, then press the Enter key. The program will ask where you want to store the expanded file. Respond by entering a full DOS pathname and filename to specify the drive, directory and name for the expanded file. For example, you may enter **C:\SURVEY\FILE1.DAT**. Note that the directory to which the file will be written (C:\SURVEY) must already exist on your hard disk. If the file (**FILE1.DAT**) already exists on the directory, you will be asked if you want to overwrite the file. Enter **Y** or **N**, as appropriate. The expanded file will be created under the filename and directory specified.

The program will ask if you want to extract specific State/Indian lands data from the survey data file. (Note: Read the file size considerations noted below before deciding how to extract the data.) To extract all of the data in the file, enter **A**. Enter **S** to extract only a subset of the data, rather than the entire file. You may select state codes from the list as instructed by the program. Note that the codes must be entered exactly as listed. After selecting the states, enter **1** to extract the file. If you make a mistake, enter **2** to re-enter the list of codes. You may enter **3** at any time to see the list of codes again, or **0** to exit the program.

2. SIZE CONSIDERATIONS

The entire expanded file for this diskette requires approximately 1.3 Megabytes of disk space. The expanded file is a standard DOS text file, with fixed-length records, one record for each house returning useable measurements. The expanded data file contains 99 ASCII text characters on each record, followed by carriage return and linefeed characters at the end of each line of text. A description of the layout of information on each record is included in the documentation for this diskette as Appendix B. The variable names listed there are the names used in EPA's analysis of the survey data.

The expanded file may be imported into a variety of DOS application programs for display and/or analysis. Most DOS applications can import DOS text files. Analysis of the data will require the use of an application program and a computer with sufficient memory available to handle a file of the required size. This should be considered when the Extract program is run. If data for all states on the disk are extracted into a single expanded file and your computer does not have additional extended or expanded memory beyond the now standard 640 Kilobytes of DOS memory, the large size of the expanded file may cause problems in many applications.

Another consideration is the number of lines (records) in the expanded file. While Excel for Windows can accommodate over 16,000 lines of data, many spreadsheet programs have a limit of approximately 8,000 lines. The entire expanded file exceeds 8,000 lines and an error will occur when importing the file into Lotus 123, for example, although sufficient memory may be available. <u>If these size problems are a concern for your program or computer, we recommend extracting the data for each state into a separate file.</u> The resulting expanded files for each state will be much smaller and problems due to size will be avoided.

3. ACCESSING DATA IN THE EXPANDED FILE

The expanded file is sorted by county within states, so that all records for a given county are grouped together in the file. For users without access to more powerful software, selected portions of the data may be viewed and printed using any word processing program that accepts DOS text files as input. For example, in version 5.0 of Wordperfect this is accomplished by the [Control-F5, 1, 2] keystroke sequence. Select a smaller font or use the landscape page orientation to print all 99 columns of data.

To conserve disk space, the expanded file does not include blank spaces between adjacent entries on a record, so a simple printout of the file as received may difficult to read. It is also difficult to analyze the data using a word processing program. DOS spreadsheet and database application programs may be used to reformat, graph and/or analyze the data.

The expanded file may be imported into a Lotus 123 spreadsheet, for example, using the [/File, Import, Text] keystroke sequence, if sufficient memory is available. The specific variables on each record may be parsed into individual numeric and label cells using the [/Data, Parse, Format, Create] keystroke sequence to specify the columns with the desired information. Then set the Input and Output ranges from the data parse menu, followed by Go. Other spreadsheet and database packages have specific procedures for importing DOS text file specified in the user reference manual.

4. CONSIDERATIONS FOR DATA ANALYSIS

This file reports short-term screening level radon measurements, conducted in accordance with prevailing EPA protocols in effect in the year of the survey. The file contains one record for each surveyed home with a useable radon measurement collected during the survey. Some data fields may have missing entries on certain records. Although attempts were made to gather complete information on each useable radon test, it was not possible to complete all items for all surveyed homes. Missing data items are indicated by a blank data field or by a single period in the data field.

The radon concentrations were estimated using a laboratory counting procedure on the exposed charcoal canisters, with a correction made for counts due to background radiation. This correction results in negative estimates of the radon concentration in some homes. These negative numbers should be considered a result of measurement error. In reality, radon concentrations are always non-negative.

The percent error variable recorded on the data file is the percentage measurement error reported by the EPA laboratory. This 2-sigma error bound was calculated based on the expected counting errors involved in the measurement process. No percentage measurement errors were reported by the laboratory for radon activities less than about 0.50 pCi/L. In the database the percent error variable is set to 0.0 on these records. For this variable, a percent error value of 0.0 should be treated as a missing value. In reality, the percentage measurement error associated with these measurements is very large.

The two problems noted above both derive from the lack of a specified Lower Limit of Detection (LLD) for the state survey data. One solution to both problems is to use the percent error variable to define the LLD for the radon activity variable. If the percent error is 0.0 and the radon activity is 0.5 pCi/L or less, then the radon activity measurement is below the LLD for the laboratory and its actual numeric value is meaningless. Alternatively, the negative activity values may be set to a small non-negative number, such as 0.05 pCi/L. This alternative method was used to calculate the survey statistics reported in this documentation.

APPENDIX B

Record Layout for State Residential Radon Surveys

Record Layout for State Residential Radon Surveys

Variable	Position	Type	Length	Description
STATE	1-2	A	2	State Postal Abbreviation (R5, R6, R7, RB, RC, RN are Indian Nations)
STATE2	3-4	A	2	State Postal Abbreviation for Indian Land Surveys (STATE = STATE2 for all other records)
STFIPS	5-6	N	2	State FIPS Code
ZIP	7-11	A	5	Zip Code
REGION	12-13	N	2	Analysis Region Code
TYPEBLDG	14	N	1	Type of Building 0 = unknown 1 = single family 2 = multi-family 3 = business 4 = school 5 = other
FLOOR	15	N	1	Floor Level 0 = basement 1 = first floor 2 = second floor or above 9 = unknown
ROOM	16	N	1	Type of Room 0 = unknown 1 = bedroom 2 = family room 3 = living room 4 = unfinished basement 5 = office 6 = classroom 7 = other

Record Layout for State Residential Radon Surveys - continued

Variable	Position	Type	Length	Description
BASEMENT	17	A	1	Is There a Basement in the Building? blank = unknown Y = Yes N = No
WINDOOR	18	A	1	House Closed or Open During Test blank = unknown O = Open C = Closed
REP	19-20	N	2	Replicate Number
STRATUM	21-22	N	2	Stratum Number
WAVE	23-25	N	3	Wave Number
STARTTM	26-29	N	4	Start Time of Test (HHMM)
STOPTM	30-33	N	4	Stop Time of Test (HHMM)
STARTDT	34-39	N	6	Start Date of Test (MMDDYY)
STOPDT	40-45	N	6	Stop Date of Test (MMDDYY)
ACTIVITY	46-53	N	8.1	Activity (pCi/L)
PCTERR	54-61	N	8.1	Percent Error (2-sigma)
ADJWT	62-74	N	13.6	Analysis Weight
DUPFLAG	75	N	1	Duplicate Flag 0 = activity from single canister 1 = average activity from duplicate canisters
ZIPFLAG	76	N	1	Flag for Zip Code (ZIP) 0 = believed accurate 1 = questionable

Record Layout for State Residential Radon Surveys - continued

Variable	Position	Type	Length	Description
CNTYFIPS	77-79	N	3	County FIPS Code
COUNTY	80-99	A	20	County Name

APPENDIX C

Description of Sample Allocation Used for Each State

ARIZONA (04)

Allocation #2 was used.
Expected DEFF = 1.041

Stratum	Geological Classification Expected Radon Level	Canisters	Relative Sampling Rates
1	AZ01 (H)	369	2.0 x
2	AZ01 (H)	1,881	1.0 x
	Total:	2,250	

INDIANA (18)

Allocation #2 was used.
Expected DEFF = 1.22

Stratum	Geological Classification Expected Radon Level	Canisters	Relative Sampling Rates
1	IN01 (H)	66	4.0 x
2	IN02 (M^+)	622	3.0 x
3	IN03 (M^-)	924	2.0 x
4	IN04 (L^+) & IN05 (L^-)	581	1.0 x
	Total:	2,193	

MASSACHUSETTS (25)

Allocation #2 was used.
Expected DEFF = 1.0

Stratum	Geological Classification Expected Radon Level	Canisters	Relative Sampling Rates
1	MA01 (H)	226	1.0 x
2	MA02 (M^+)	873	1.0 x
3	MA03 (M^-)	894	1.0 x
4	MA04 (L)	7	1.0 x
	Total:	2,000	

MINNESOTA (27)

Allocation #3 was used.
Expected DEFF = 1.1475

Stratum	Geological Classification Expected Radon Level	Canisters	Relative Sampling Rates
1	MN01 (H) & MN02 (M)	22	4.0 x
2	MN01 (H) & MN03 (L)	283	2.5 x
3	MN02 (M)	329	2.0 x
4	MN03 (L)	146	1.5 x
5	MN01 (H)	220	1.0 x
	Total:	1,000	

MISSOURI (29)

Allocation #4 was used.
Expected DEFF = 1.164

Stratum	Geological Classification Expected Radon Level	Canisters	Relative Sampling Rates
1	MO01 (H)	341	3.0 x
2	MO01 (H), MO02 (M^+), MO03 (M^-), MO04 (L^+), MO05 (L^-)	1,151	2.0 x
3	MO03 (M^-), MO04 (L^+), MO05 (L^-)	758	1.0 x
	Total:	2,250	

NORTH DAKOTA (38)

Allocation #3 was used.
Expected DEFF = 1.24

Stratum	Geological Classification Expected Radon Level	Canisters	Relative Sampling Rates
1	ND01 (H)	199	4.2 x
2	ND01 (H)	333	4.1 x
3	ND02 (M)	654	2.1 x
4	ND02 (M), ND03 (L)	411	1.9 x
5	ND03 (L)	403	1.0 x
	Total:	2,000	

PENNSYLVANIA (42)

Allocation #3 was used.
Expected DEFF - 1.289

Reporting Group	Canisters	Relative Sampling Rates
1	317	1.0 x
2	289	2.0 x
3	271	2.0 x
4	347	7.0 x
5	276	2.0 x
6	291	3.0 x
7	257	7.0 x
8	255	4.0 x
9	328	2.0 x
10	369	2.0 x
Total:	3,000	

Region 5 Indians (26, 27, 55)

Allocation #3 was used.

State	Stratum	Canisters	
MI	1	60	
MI	2	12	
MI	3	54	
MI	4	75	
	Total:		211
MN	5	79	
MN	6	165	
MN	7	46	
	Total:		290
WI	8	258	
WI	9	111	
WI	10	117	
WI	11	13	
			500
	Total:	1,000	

Table C-1 Distribution of Canisters per County for Arizona

COUNTY	REGION	# CANISTERS
APACHE	1	15
COCHISE	3	39
COCONINO	1	89
GILA	2	13
GRAHAM	2	29
GREENLEE	2	8
LA PAZ	3	2
MARICOPA	3	765
MOHAVE	2	99
NAVAJO	1	57
PIMA	3	260
PINAL	3	33
SANTA CRUZ	3	13
YAVAPAI	2	51
YUMA	3	34

Table C-1 Distribution of Canisters per County for Indiana

COUNTY	REGION	# CANISTERS
ADAMS	2	14
ALLEN	2	169
BARTHOLOMEW	4	28
BENTON	1	2
BLACKFORD	3	4
BOONE	3	9
BROWN	3	3
CARROLL	1	7
CASS	2	6
CLARK	4	92
CLAY	5	8
CLINTON	3	7
CRAWFORD	5	2
DAVIESS	5	5
DE KALB	2	21
DEARBORN	3	6
DECATUR	3	5
DELAWARE	3	16
DUBOIS	5	5
ELKHART	2	76
FAYETTE	3	6
FLOYD	4	32
FOUNTAIN	3	13
FRANKLIN	3	4
FULTON	2	9
GIBSON	5	16
GRANT	3	13
GREENE	5	16
HAMILTON	3	23
HANCOCK	3	8
HARRISON	4	19
HENDRICKS	3	22
HENRY	3	11
HOWARD	3	22
HUNTINGTON	2	13
JACKSON	4	7
JASPER	1	11
JAY	3	5
JEFFERSON	4	16
JENNINGS	4	19
JOHNSON	3	34
KNOX	5	9
KOSCIUSKO	2	30
LA PORTE	1	66
LAGRANGE	2	9

Table C-1 Distribution of Canisters per County for Indiana (Continued)

COUNTY	REGION	# CANISTERS
LAKE	1	125
LAWRENCE	4	28
MADISON	3	27
MARION	3	115
MARSHALL	2	3
MARTIN	5	5
MIAMI	2	28
MONROE	4	30
MONTGOMERY	3	21
MORGAN	3	7
NEWTON	1	12
NOBLE	2	20
OHIO	3	4
ORANGE	4	11
OWEN	3	5
PARKE	5	7
PERRY	5	3
PIKE	5	8
PORTER	1	84
POSEY	5	6
PULASKI	1	5
PUTNAM	3	6
RANDOLPH	3	9
RIPLEY	3	6
RUSH	3	1
SCOTT	4	21
SHELBY	3	7
SPENCER	5	11
STARKE	1	8
STEUBEN	2	13
ST. JOSEPH	1	114
SULLIVAN	5	12
SWITZERLAND	3	2
TIPPECANOE	1	39
TIPTON	3	5
UNION	0	0
VANDERBURGH	5	32
VERMILLION	5	8
VIGO	5	34
WABASH	2	15
WARREN	3	4
WARRICK	5	21
WASHINGTON	4	10
WAYNE	3	18
WELLS	2	7
WHITE	1	16
WHITLEY	2	23

Table C-1 Distribution of Canisters per County for Massachusetts

COUNTY	REGION	# CANISTERS
BARNS	1	97
BERKSHIRE	3	47
BRISTOL	9	114
DUKES	1	6
ESSEX	7	201
FRANKLIN	2	27
HAMPDEN	4	125
HAMPSHIRE	2	54
MIDDLESEX	6	391
NANTUCKET	0	0
NORFOLK	8	162
PLYMOUTH	10	141
SUFFOLK	11	75
WORCESTER	5	219

Table C-1 Distribution of Canisters per County for Minnesota

COUNTY	REGION	# CANISTERS
AITKIN	5	4
ANOKA	3	52
BECKER	2	3
BELTRAMI	2	7
BENTON	5	4
BIG STONE	5	3
BLUE EARTH	4	14
BROWN	4	4
CARLTON	5	10
CARVER	5	6
CASS	2	5
CHIPPEWA	5	4
CHISAGO	5	6
CLAY	2	14
CLEARWATER	2	4
COOK	1	2
COTTONWOOD	5	4
CROW WING	2	12
DAKOTA	3	63
DODGE	4	3
DOUGLAS	5	9
FARIBAULT	4	6
FILLMORE	4	2
FREEBORN	4	9
GOODHUE	4	14
GRANT	0	0
HENNEPIN	3	105
HOUSTON	4	6
HUBBARD	2	5
ISANTI	5	3
ITASCA	2	11
JACKSON	5	5
KANABEC	5	4
KANDIYOHI	5	4
KITTSON	2	3
KOOCHICHING	2	7
LAC QUI PARLE	5	2
LAKE	1	9
LAKE OF THE WOODS	2	4
LE SUEUR	4	5
LINCOLN	5	4
LYON	5	8
MAHNOMEN	2	1
MARSHALL	2	9
MARTIN	5	7

Table C-1 Distribution of Canisters per County for Minnesota (Continued)

COUNTY	REGION	# CANISTERS
MCLEOD	5	13
MEEKER	5	5
MILLE LACS	5	2
MORRISON	2	9
MOWER	4	13
MURRAY	5	1
NICOLLET	5	4
NOBLES	5	3
NORMAN	2	3
OLMSTED	4	23
OTTER TAIL	2	8
PENNINGTON	2	3
PINE	5	6
PIPESTONE	5	4
POLK	2	4
POPE	5	2
RAMSEY	3	32
RED LAKE	0	0
REDWOOD	5	5
RENVILLE	5	3
RICE	4	11
ROCK	5	2
ROSEAU	2	14
SCOTT	5	13
SHERBURNE	5	8
SIBLEY	5	4
STEARNS	2	25
STEELE	4	10
STEVENS	5	2
ST. LOUIS	1	116
SWIFT	5	4
TODD	2	3
TRAVERSE	5	4
WABASHA	4	7
WADENA	2	5
WASECA	4	4
WASHINGTON	3	46
WATONWAN	5	3
WILKIN	2	1
WINONA	4	13
WRIGHT	5	13
YELLOW MEDICINE	5	2

Table C-1 Distribution of Canisters per County for Missouri

COUNTY	REGION	# CANISTERS
ADAIR	2	6
ANDREW	1	6
ATCHISON	1	4
AUDRAIN	3	6
BARRY	5	10
BARTON	5	10
BATES	5	11
BENTON	5	7
BOLLINGER	6	10
BOONE	3	13
BUCHANAN	1	39
BUTLER	6	12
CALDWELL	1	3
CALLAWAY	3	3
CAMDEN	3	24
CAPE GIRARDEAU	6	10
CARROLL	1	4
CARTER	6	2
CASS	1	57
CEDAR	5	8
CHARITON	2	4
CHRISTIAN	5	7
CLARK	2	5
CLAY	1	94
CLINTON	1	4
COLE	3	34
COOPER	3	4
CRAWFORD	3	7
DADE	5	1
DALLAS	5	4
DAVIESS	1	7
DE KALB	1	4
DENT	3	6
DOUGLAS	6	8
DUNKLIN	6	12
FRANKLIN	4	40
GASCONADE	3	7
GENTRY	1	5
GREENE	5	42
GRUNDY	2	4
HARRISON	1	6
HENRY	5	28
HICKORY	5	2
HOLT	1	2
HOWARD	3	10

Table C-1 Distribution of Canisters per County for Missouri (Continued)

COUNTY	REGION	# CANISTERS
HOWELL	6	15
IRON	6	7
JACKSON	1	271
JASPER	5	40
JEFFERSON	4	49
JOHNSON	1	32
KNOX	2	3
LACLEDE	3	12
LAFAYETTE	1	34
LAWRENCE	5	11
LEWIS	2	1
LINCOLN	3	6
LINN	2	5
LIVINGSTON	2	2
MACON	2	5
MADISON	6	7
MARIES	3	2
MARION	2	8
MCDONALD	5	5
MERCER	2	2
MILLER	3	10
MISSISSIPPI	6	3
MONITEAU	3	4
MONROE	2	7
MONTGOMERY	3	2
MORGAN	3	8
NEW MADRID	6	7
NEWTON	5	11
NODAWAY	1	8
OREGON	6	3
OSAGE	3	6
OZARK	6	5
PEMISCOT	6	1
PERRY	6	4
PETTIS	3	18
PHELPS	3	14
PIKE	2	9
PLATTE	1	25
POLK	5	15
PULASKI	3	12
PUTNAM	2	3
RALLS	2	1
RANDOLPH	2	6
RAY	1	17
REYNOLDS	6	1

Table C-1 Distribution of Canisters per County for Missouri (Continued)

COUNTY	REGION	# CANISTERS
RIPLEY	6	11
SALINE	2	9
SCHUYLER	2	1
SCOTLAND	2	4
SCOTT	6	13
SHANNON	6	6
SHELBY	2	3
STE. GENEVIEVE	6	11
STODDARD	6	9
STONE	4	9
ST. CHARLES	5	50
ST. CLAIR	5	3
ST. FRANCOIS	6	48
ST. LOUIS	4	91
ST. LOUIS CITY	4	207
SULLIVAN	2	2
TANEY	5	10
TEXAS	6	8
VERNON	5	13
WARREN	3	6
WASHINGTON	3	4
WAYNE	6	8
WEBSTER	5	15
WORTH	1	2
WRIGHT	6	7

Table C-1 Distribution of Canisters per County for North Dakota

COUNTY	REGION	# CANISTERS
ADAMS	1	23
BARNES	4	38
BENSON	6	8
BILLINGS	1	9
BOTTINEAU	3	33
BOWMAN	1	31
BURKE	5	4
BURLEIGH	5	101
CASS	4	171
CAVALIER	6	14
DICKEY	6	11
DIVIDE	5	4
DUNN	2	28
EDDY	6	4
EMMONS	2	15
FOSTER	6	7
GOLDEN VALLEY	2	7
GRAND FORKS	4	172
GRANT	1	23
GRIGGS	6	10
HETTINGER	1	31
KIDDER	3	8
LA MOURE	6	5
LOGAN	3	13
MCHENRY	3	30
MCINTOSH	5	9
MCKENZIE	2	6
MCLEAN	2	17
MERCER	2	46
MORTON	1	99
MOUNTRAIL	5	20
NELSON	6	26
OLIVER	1	19
PEMBINA	1	59
PIERCE	3	17
RAMSEY	6	18
RANSOM	6	7
RENVILLE	5	9
RICHLAND	4	46
ROLETTE	3	20
SARGENT	4	10
SHERIDAN	5	5
SIOUX	2	2
SLOPE	1	7
STARK	1	122

Table C-1 Distribution of Canisters per County for North Dakota (Continued)

COUNTY	REGION	# CANISTERS
STEELE	4	7
STUTSMAN	6	40
TOWNER	3	10
TRAILL	4	26
WALSH	6	49
WARD	5	66
WELLS	6	11
WILLIAMS	5	23

Table C-1 Distribution of Canisters per County Pennsylvania

COUNTY	REGION	# CANISTERS
ADAMS	3	416
ALLEGHENY	9	4
ARMSTRONG	5	1
BEAVER	4	1
BEDFORD	3	14
BERKS	1	40
BLAIR	3	32
BRADFORD	7	40
BUCKS	1	46
BUTLER	4	97
CAMBRIA	5	42
CAMERON	6	5
CARBON	3	18
CENTRE	3	55
CHESTER	1	4
CLARION	6	13
CLEARFIELD	6	28
CLINTON	3	7
COLUMBIA	3	10
CRAWFORD	6	21
CUMBERLAND	2	45
DAUPHIN	2	87
DELAWARE	1	1
ELK	6	17
ERIE	6	70
FAYETTE	5	28
FOREST	6	3
FRANKLIN	3	20
FULTON	3	3
GREENE	5	7
HUNTINGDON	3	7
INDIANA	5	17
JEFFERSON	6	14
JUNIATA	3	84
LACKAWANNA	8	2
LANCASTER	2	69
LAWRENCE	4	56
LEBANON	2	21
LEHIGH	1	23
LUZERNE	8	118
LYCOMING	3	28
MCKEAN	6	64
MERCER	6	2
MIFFLIN	3	10
MONROE	7	111

Table C-1 Distribution of Canisters per County for Pennsylvania (Continued)

COUNTY	REGION	# CANISTERS
MONTGOMERY	1	1
MONTOUR	3	6
NORTHAMPTON	1	26
NORTHUMBERLAND	3	15
PERRY	2	6
PHILADELPHIA	10	125
PIKE	7	8
POTTER	7	12
SCHUYLKILL	3	32
SNYDER	3	28
SOMERSET	5	3
SULLIVAN	3	1
SUSQUEHANNA	7	20
TIOGA	7	29
UNION	3	34
VENANGO	6	1
WARREN	6	15
WASHINGTON	5	45
WAYNE	7	100
WESTMORELAND	5	1
WYOMING	7	78
YORK	2	2

APPENDIX D

Regional Radon Coordinators and Sources of Information Concerning Other State-Wide Radon Studies

Regional Radon Coordinators

EPA REGION	REGIONAL OFFICE	CONTACT
1	U.S. Environmental Protection Agency John F. Kennedy Federal Building Room 2311 Boston, MA 02203	Mona Haywood (617) 565-9402
2	U.S. Environmental Protection Agency 26 Federal Plaza Room 1137-L New York, NY 10278	Lorainne Koehler (212) 264-0546
3	U.S. Environmental Protection Agency (3AM12) 841 Chestnut Street Philadelphia, PA 19107	Lewis Felleisen (215) 597-8326
4	U.S. Environmental Protection Agency 345 Courtland Street, NE Atlanta, GA 30365	Paul Wagner (404) 347-3907
5	U.S. Environmental Protection Agency Mail Code (AT-18J) 77 West Jackson Blvd. Chicago, IL 60604	Julie Beckman (312) 886-6063
6	U.S. Environmental Protection Agency Air Enforcement Branch (6T-E) 1445 Ross Avenue Dallas, TX 75202	Michael Miller (214) 655-7550
7	U.S. Environmental Protection Agency 726 Minnesota Avenue Kansas City, KS 66101	Bob Hunt (913) 551-7611
8	U.S. Environmental Protection Agency (8HWM-RP) Suite 500 999 18th Street Denver, CO 80202	Milton W. Lammering (303) 293-1440
9	U.S. Environmental Protection Agency (A1-1) 75 Hawthorne Street San Francisco, CA 94105	Louise Hill (415) 744-1046
10	U.S. Environmental Protection Agency (AT-082) 1200 Sixth Avenue Seattle, WA 98101	Misha Vakoc (206) 553-7299

Sources of Information Concerning Other State-Wide Radon Studies		
STATE	AGENCY	CONTACT
New Jersey	Department of Environmental Protection 729 Alexander Road Princeton, NJ 08540	Robert Stern (800) 648-0394 (609) 987-6402
New York	State Health Department Bureau of Environmental Radiation Protection Corning Tower Albany, NY 12237	Laurence Keefe (800) 458-1158 (518) 458-6450
North Carolina	Department of Human Resources Radiation Protection Section 701 Barbour Drive Raleigh, NC 27603-2008	Dr. Felix Fong (919) 733-4283
Idaho	Department of Health and Welfare Bureau of Preventive Medicine 450 West State Street Boise, ID 83720	Janne Mitten (208) 334-5927
Florida	Department of Health and Rehabilitative Services 1317 Winewood Boulevard Tallahassee, FL 32399-0700	N. Michael Gilly (800) 543-8279 (904) 488-1525
South Carolina	Department of Health and Environmental Control Bureau of Radiological Health 2600 Bull Street Colombia, SC 29201	Nolan Bivens (803) 734-4700
Oregon	Department of Human Services Health Division 1400 SW 5th Avenue Portland, OR 97201	Ray Paris (503) 229-5797
Washington	Department of Health Office of Radiation Protection Airdustrial Building 5, LE-13 Olympia, WA 98504	Robert Mooney (206) 586-3303

STATE	AGENCY	CONTACT
Montana	Department of Health and Environmental Sciences Cogswell Building Helena, MT 59620	Adrian Howe (406) 444-3671
New Hampshire	Division of Public Health Serv. Bureau of Radiological Health 6 Hazen Drive Concord, NH 03301	Joy Hanington (603) 271-4674
Virginia	Department of Health Bureau of Radiological Health 109 Governor Street Richmond, VA 23219	Leslie Foldesi (800) 468-0138 (804) 786-5932
Nevada	Department of Human Resources Radiological Health Section 505 East King Street, Rm. 203 Carson City, NV 89710	Stan Marshall (702) 885-5394
Louisiana	Louisiana Nuclear Energy Division Department of Environmental Qual. P.O. Box 14690 Baton Rouge, LA 70898	Jay Mason (504) 925-4518

APPENDIX E

Procedures for Estimating Weighted Means, Proportions, Standard Errors and Confidence Intervals for Indian Lands

Procedures for Estimating Weighted Means, Proportions, Standard Errors and Confidence Intervals for Indian Lands

The EPA's Region 5 Indian lands consist of 29 reservations. For purposes of the radon survey, the area was stratified according to reservation and a simple random sample of households was selected within each reservation or stratum. Formulas for generating estimates of weighted means, proportions and standard errors are given below. An approximate 95 percent confidence interval can be derived by adding to and subtracting from the estimate two standard errors of the estimate.

NOTATION

Let, Y_{hi} = observed radon measurement for the i^{th} household in stratum h ($i=1, ..., n_h$ and $h=1,...,H$);

W_{hi} = sampling weight associated with Y_{hi} ;

$$J_{rh} = \begin{cases} 1 & \text{if stratum h is included in the } r^{th} \text{ region} \\ 0 & \text{otherwise;} \end{cases}$$

$$I_{xhi} = \begin{cases} 1 & \text{if measurement on } i^{th} \text{ household in stratum h is greater than X pCi/L} \\ 0 & \text{otherwise;} \end{cases}$$

n_h = number of sample households in stratum h;
H = number of strata;

$$N_h = \sum_{i=1}^{n_h} W_{hi} \quad ;$$

$$N_r = \sum_{h=1}^{H} J_{rh} \quad N_h \quad ;$$

$$N = \sum_{h=1}^{H} N_h \quad ;$$

$$s^2_h = \frac{\sum_{i=1}^{n_h} Y^2_{hi} - \left| \sum_{i=1}^{n_h} Y_{hi} \right|^2 / n_h}{n_h - 1}; \quad \text{and}$$

$$S.E.(est.) = [Var(est.)]^{1/2}.$$

ESTIMATION:

The true mean radon level for the h^{th} stratum or reservation can be estimated as

$$Y^*_h = \frac{\sum_{i=1}^{n_h} W_{hi} \, Y_{hi}}{\sum_{i=1}^{n_h} W_{hi}} \tag{1}$$

The mean radon level for the r^{th} region, consisting of two or more reservations, is given by the weighted average of the strata making up the region, namely

$$Y^*_r = \frac{\sum_{h=1}^{H} J_{rh} \, N_h \, Y^*_h}{N_r} \tag{2}$$

The variance of Y^*_r is estimated as

$$Var(Y^*_r) = \frac{1}{N^2_r} \sum_{h=1}^{H} J_{rh} \, N_h(N_h - n_h) \left| \frac{s^2_h}{n_h} \right|, \tag{3}$$

and the standard error is obtained as $S.E.(Y^*_r) = [Var(Y^*_r)]^{1/2}$. A weighted average of all strata means provides an estimate of the overall mean,

$$Y^* = \frac{\sum_{h=1}^{H} N_h \, Y^*_h}{N} \tag{4}$$

The variance of Y^* is estimated as

$$Var(Y^*) = \frac{1}{N^2} \sum_{h=1}^{H} N_h (N_h - n_h) \left| \frac{s^2_h}{n_h} \right| , \tag{5}$$

and the standard error is obtained as $s.e.(Y^*) = [Var(Y^*)]^{1/2}$.

The true proportion of households in the h^{th} stratum with radon levels exceeding X pCi/L can be estimated as

$$P^*_h = \frac{\sum_{h=1}^{H} W_{hi} \, I_{xhi}}{\sum_{i=1}^{n_h} W_{hi}} \tag{6}$$

The proportion of households in the r^{th} region (i.e., combination of reservations) with radon levels exceeding X pCi/L is given by

$$P^*_r = \frac{\sum_{h=1}^{H} J_{rh} \, N_h \, P^*_h}{N_r} \tag{7}$$

The variance of P^*_r is estimated as

$$Var(P^*_r) = \frac{1}{N^2_r} \sum_{h=1}^{H} J_{rh} \, N_h (N_h - n_h) \left| \frac{P^*_h (1 - P^*_h)}{n_h - 1} \right| , \tag{8}$$

and the standard error is obtained as $s.e.(P^*_r) = [Var(P^*_r)]^{1/2}$. A weighted average of all strata proportion provides an overall proportion, namely

$$P^* = \frac{\sum_{h=1}^{H} N_h P^*_h}{N} . \qquad (9)$$

The estimated variance of P^* is given by

$$Var(P^*) = \frac{1}{N^2} \sum_{h=1}^{H} N_h(N_h - n_h) \left| \frac{P^*_h(1-P^*_h)}{n_h - 1} \right| , \qquad (10)$$

and the standard error is obtained as $s.e.(P^*) = [Var(P^*)]^{1/2}$.